

# Modelling habitat preference and estimating the spatial distribution of Australian Sea Lions (*Neophoca cinerea*); “A first exploration”

G. M. Aarts & S. Brasseur

Report number C107/08



Institute for Marine Resources and Ecosystem Studies

Wageningen **IMARES**

Location Texel

Client: South Australian Research & Development Institute (SARDI) - Aquatic Sciences A/Prof  
Simon Goldsworthy  
SARDI Aquatic Sciences  
PO Box 120, Henley Beach, SA 5022, AUSTRALIA

Publication Date: December 2008

- Wageningen **IMARES** conducts research providing knowledge necessary for the protection, harvest and usage of marine and coastal areas.
- Wageningen **IMARES** is a knowledge and research partner for governmental authorities, private industry and social organisations for which marine habitat and resources are of interest.
- Wageningen **IMARES** provides strategic and applied ecological investigation related to ecological and economic developments.

© 2008 Wageningen **IMARES**

Wageningen IMARES is a cooperative research organisation formed by Wageningen UR en TNO. We are registered in the Dutch trade record Amsterdam nr. 34135929, BTW nr. NL 811383696B04.

The Management of IMARES is not responsible for resulting damage, as well as for damage resulting from the application of results or research obtained by IMARES, its clients or any claims related to the application of information found within its research. This report has been made on the request of the client and is wholly the client's property. This report may not be reproduced and/or published partially or in its entirety without the express written consent of the client.



A\_4\_3\_2V4

# Contents

Contents .....	3
Abstract.....	4
1. Introduction.....	5
2. Methods .....	5
2.1 Data pre-processing.....	5
2.2 Accounting for unequal habitat availability; defining the response variable.....	9
2.3 Modelling preference .....	9
2.4 Predicting the spatial distribution .....	10
3. Results .....	11
3.1 The effect of environmental variables on preference .....	11
3.2 Spatial prediction of usage .....	12
3.3 Individual difference in habitat preference .....	13
4. Discussion .....	16
4.1 Spatial distribution and habitat preference of the Australian sea lion.....	16
4.2 Incorporating the right environmental variables .....	16
Acknowledgements.....	17
References .....	17
Referees and Authors .....	18

## Abstract

Managing the Australian sea lion (*Neophoca cinerea*) population and mitigating its interactions with commercial fisheries, requires an understanding of their spatial distribution and habitat preference at sea. Numerous wildlife telemetry devices have been attached to individual seals from different colonies, providing a detailed insight into their movement and activities. However, as data are only available from some individuals from 16 of 40 colonies in South Australia, these data represents only a small proportion of the population. Moreover, some colonies are poorly represented. To estimate the spatial distribution of the entire South Australian population, one can first investigate why individuals visit certain places and use this information to predict the spatial distribution for other regions lacking data.

In this study we fit Generalized Additive Mixed Models (GAMMs) to wildlife telemetry data collected from adult female Australian sea lions to investigate the species' habitat preference for the variables distance to the colony, depth and slope. The results show that in general they have a higher preference for shallow areas, places close to the colony and a steep slope, but they also display large individual variability. Preference for these variables does not seem to differ between individuals of different sizes. In some of the most western colonies, some individuals seem to spend more time further away from the colony, but apart from this, there is no apparent effect of the longitudinal capture position on preference.

The predicted spatial distribution is largely driven by the distance from the departure colony and the actual distribution of individuals among the different colonies. Although depth and slope does seem to effect their distribution to some extent, the most important variables that explain fine-scale foraging activities at-sea are probably missing. Future studies using GPS transmitters attached to animals, that yield high resolution locations and more detailed environmental data, in combination with the analytical technique presented here, should provide more insight there foraging decisions. This should eventually also improve the spatial prediction of the population as a whole.

# 1. Introduction

The total population of Australian sea lions is currently estimated to be about 14,000. They are protected under the Commonwealth Environment Protection and Biodiversity Conservation Act, and were listed as Threatened in February 2005. The geographical overlap between their foraging space, and commercial fisheries has lead to increased risk of bycatch, especially from demersal gillnet fisheries (Goldsworthy and Page 2007). The breeding asynchrony between small and closed sub-populations makes them particularly sensitive to such additional mortality (Goldsworthy and Page 2007).

Managing the species and mitigating its interactions with the fisheries, requires an understanding of the overlap between the sea lions' spatial distribution of fishing activities. Technical advances in the last decades has enabled scientist to remotely follow individual animals using wildlife telemetry devices. The Australian sea lion project has made full use of these advantages and has deployed quite a few individuals with these tags. However, due to logistic and financial constraints, still only a relatively small proportion of the population has been tracked. Consequently, many populations have been poorly sampled, or not sampled at all.. This lack of knowledge may hamper future conservation and management of the species and fishery.

The spatial distribution of any organism is largely influenced by the availability and distribution of resources and the restrictions the environment imposes on their activities. Quantifying the link between the distribution of the species and different environmental variables can help scientists to understand how and why animals visit certain places (Boyce and McDonald 1999). This relationship can then be used to predict the distribution of the organism elsewhere (Guisan and Zimmermann 2000, Aarts et al. 2008)

In the case of the Australian sea lion, their distribution and abundance at sea is largely influenced by the number of individuals that make use of the colonies in the vicinity. Data on land-based counts are available and can therefore be used to inform their distribution and abundance at sea (Matthiopoulos et al. 2004).

The main objective of this study was to first model the relationship between the individual tracked sea lions and the available environmental variables depth, slope and distance to the colony from which forage trips are made. Finally, using the count data on the colonies, estimate of the distribution of the whole population at sea according to the methods described in (Aarts et al. 2008)

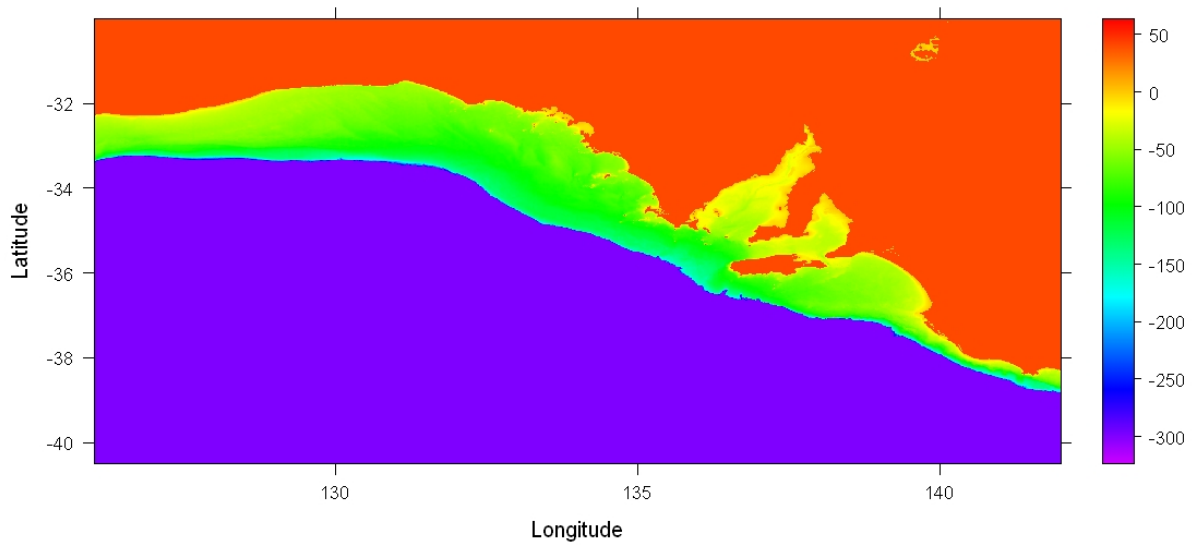
## 2. Methods

### 2.1 Data pre-processing

#### *Environmental data*

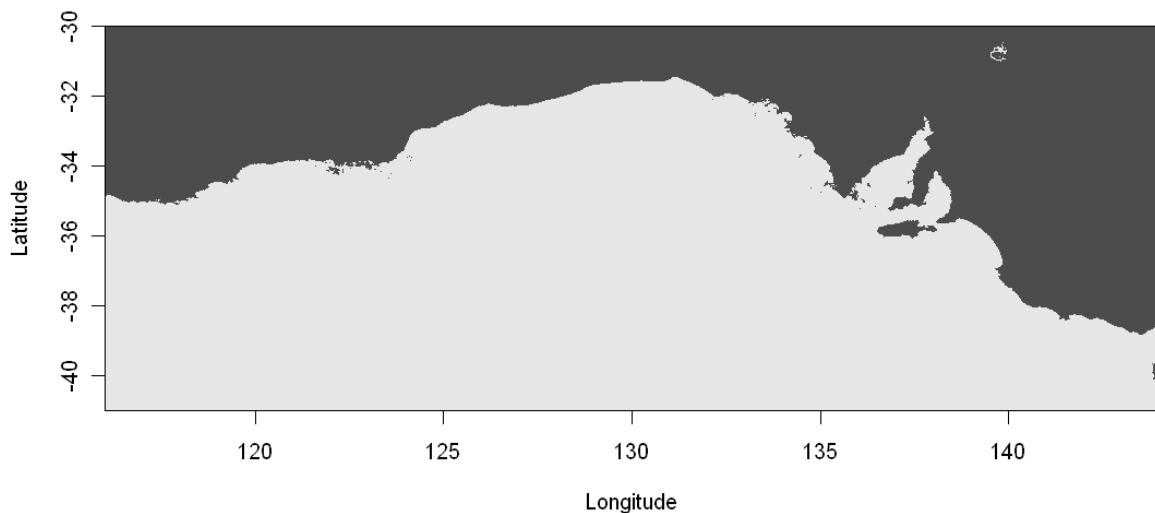
Two environmental data sources were available; depth and slope. The depth grid was used to generate a grid of land (depths > 0 m) and sea (depths < 0 m), denoted by 1 and 0 respectively. The land-sea grid was then used to estimate the shortest at-sea distance to each colony. A more detailed description of data pre-processing is provided below.

The baseline data for depth consists of two files; 'Top left.xyz2.txt' and 'top middle.xyz2.txt'. Combining these tables produced a regular 0.01° by 0.01° grid for the region (Fig. 1).

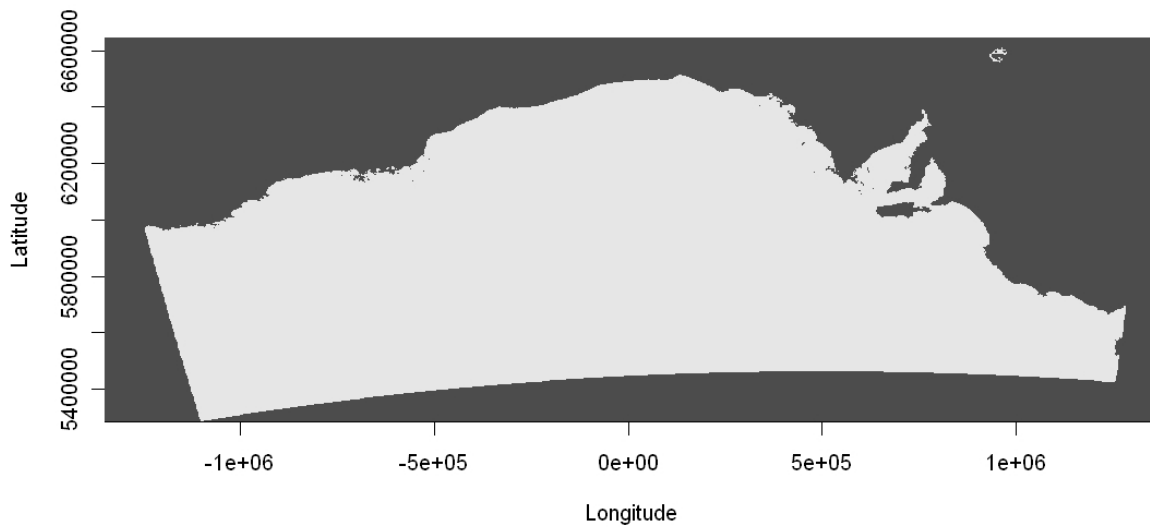


**Figure 1.** Depth map of southern Australian. The area off the continental shelf which may be as deep as 7000m, is truncated to 300m to allow for a better visual display.

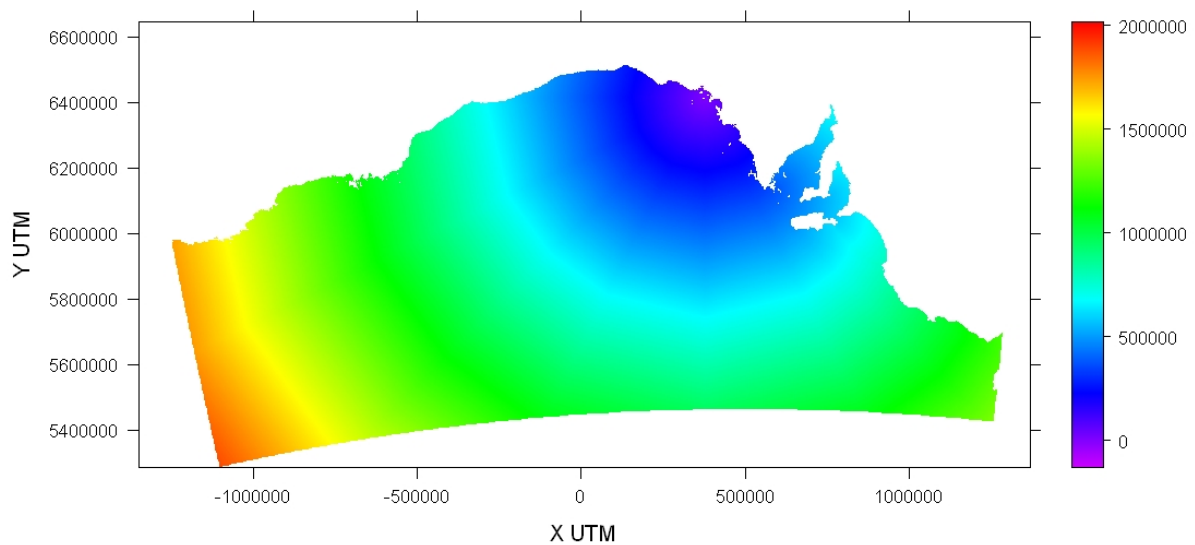
To calculate the distance to the colony, one could calculate for each point in space the straight-line distance to the colony in question. However this does not account for obstacles that prevent seals from moving over land. By using a self-writing R-routine, it is possible to estimate the at-sea distance. The exact procedure is as follows: first a land-sea grid in geographical coordinates is created by defining depths  $> 0$  as land and depths  $< 0$  as sea (Fig. 2): next, a 2 by 2km point grid in UTM53S projection is created and the corresponding land sea value extracted for each point (Fig. 3). By stepping from cell to cell, avoiding cells that are on land, it is possible to calculate the shortest at-sea distance (Fig. 4). Such at-sea distance maps are generated for each colony.



**Figure 2.** Land-sea grid in geographical coordinates



**Figure 3.** Land-sea grid in UTM53S projection. .



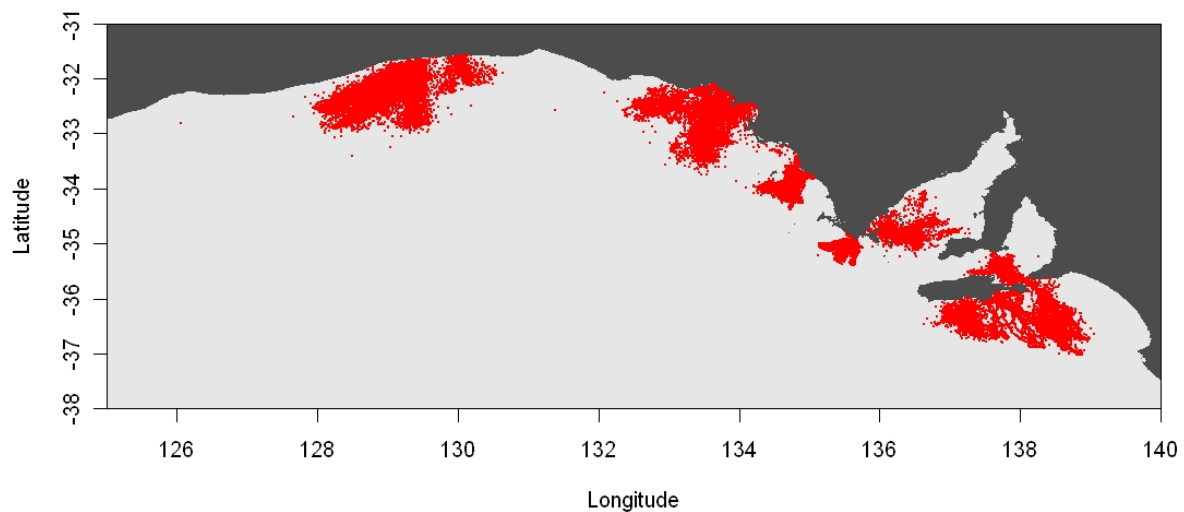
**Figure 4.** Map (UTM53S projection) of at-sea distance (in meters) relative to one colony.

Data on the slope was based on "Sea\_lion\_grid\_clipped\_bathymetry.csv", which had an approximately regular grid (0.01 by 0.009 degree latitude and longitude, respectively). For each depth point less than 500m depth, the nearest slope value was extracted. Similar to depth and land-sea, the data was stored as an ascii raster grid.

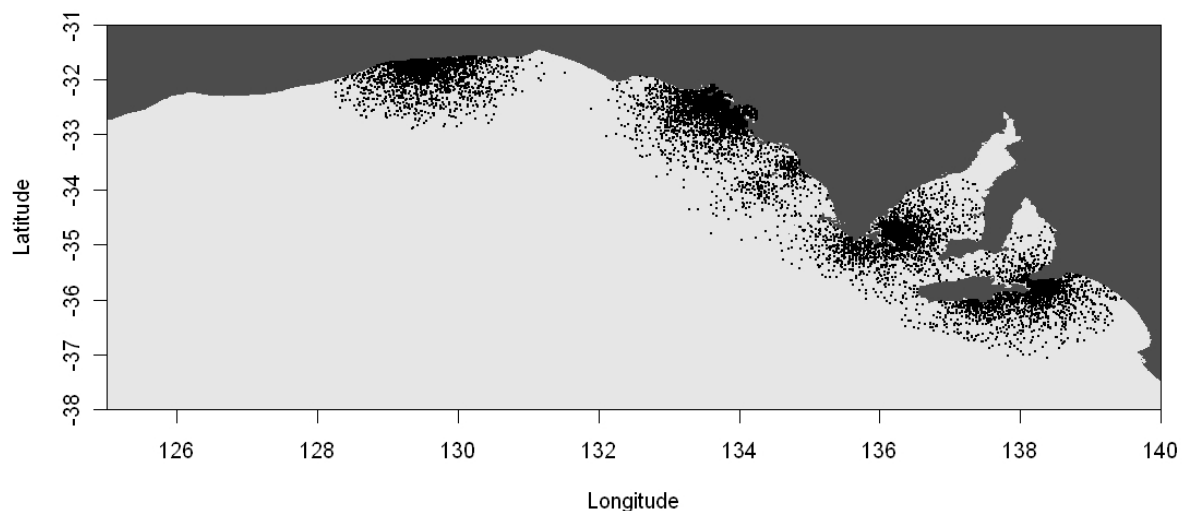
#### *Satellite telemetry data*

In total there were 89,742 filtered locations, from 196 individuals. From these locations, trips were defined. A trip started when the distance to the nearest colony exceeded 5km. Using this criteria, for each location the colony from where the trip started was defined (start colony). 112 locations that were not associated to a colony were removed from the analysis. For each location the corresponding depth, slope and distance to the start colony was extracted. Locations missing data on any environmental variables were also removed. For computational reasons it was not possible to use all data points. Therefore a sample was

taken. The animal locations used should reflect of the proportion of time animals spend in certain environmental conditions. Due to differences in behaviour and satellite reception, this assumption may be violated. By increasing the probability of selecting animal locations at times with relative few other telemetry locations, some of this bias will be reduced. This can be done by selecting animal locations with a probability proportional to the mean time to the previous and next location. In this model only data on adult females (age-sex class = 1) were used, representing 80 % of all data (Fig. 5).



**Figure 5.** All wildlife telemetry locations from adult female Australian sea lions.



**Figure 6.** The pseudo-absences selected based on a null-model of usage (see section 2.2) that together with the animal observations (Fig. 5) define the response variable.



### Colonies

Data on colonies size and foraging effort were stored in "ASL population based seal foraging day.csv". For each age and sex the foraging effort in seal days was calculated. This was the sum of the number of days that sea lions forage from the colony in question, combined for all individuals. In the remainder of the text this will be denoted as foraging effort. This table does not contain all colonies that are used by the tagged Australian sea lions, in contrast to the table "colonies\_names.csv". Therefore the "colonies\_names.csv" table was used as a reference and for each colony in that table the effort was extracted from the "ASL population based seal foraging day.csv" table.

## 2.2 Accounting for unequal habitat availability; defining the response variable

In a first attempt to understand why animals visit certain places, one could extract the underlying environmental conditions for each animal location and look at how many animal locations occur at which values for the environmental variable. This is called niche based modelling. However the relationship between animal locations and the environmental variables is not only influenced by the preference of the individuals, but also by the relative availability of different habitats in the study region. Irrespective of preference, the animal will be observed more in those habitats that are more frequently available in the study area. So if we assume that preference does not change with changes in the absolute availability of habitats, preference is the amount of usage of a habitat divided by its availability. Although this may be an unreasonable assumption, it is in fact implicitly assumed in all habitat analysis studies (Myrsetrud and Ims 1998).

To account for unequal habitat availability, the animal observations are augmented with a set of pseudo-absences, reflecting the conditions that are available to each animal (Keating and Cherry 2004, Johnson et al. 2006, Lele and Keim 2006, Aarts et al. 2008). These pseudo-absences can be uniformly selected from space or can be based on some null-model of usage. The latter is often more computationally efficient (Aarts et al. 2008). In this study we first modelled spatial distribution as a function of distance, then treated this model as the null-model of usage (Matthiopoulos 2003) and generate pseudo-absences based on these model-predictions. Fig. 6 shows the distribution of those pseudo-absences

## 2.3 Modelling preference

The presences and pseudo-absences were modelled as non-linear functions (Hastie and Tibshirani 1990) of the environmental variables depth, slope and distance to the colony, using Generalized Additive Mixed-effect Models (GAMMs) (Wood 2006, Aarts et al. 2008).

$$\begin{aligned}\hat{u}_i &\sim B(1, h_i) \\ h_i &= g^{-1}(\eta_i) = \frac{e^{\eta_i}}{1 + e^{\eta_i}} \\ \eta_i &= b_0 + s_1(\text{distance}_i) + s_2(\text{depth}_i) + s_3(\text{slope}_i)\end{aligned}\tag{1}$$

Where  $\hat{u}_i$  is the response variable which is either a one (animal location) or zero (pseudo absence) which is modelled assuming a binomial ( $B$ ) distribution. The logit ( $g$ ) of this mean probability is modelled as a function of the linear predictor  $\eta$  which contains smooth functions of the three environmental variables; distance to the colony, depth and slope.

## 2.4 Predicting the spatial distribution

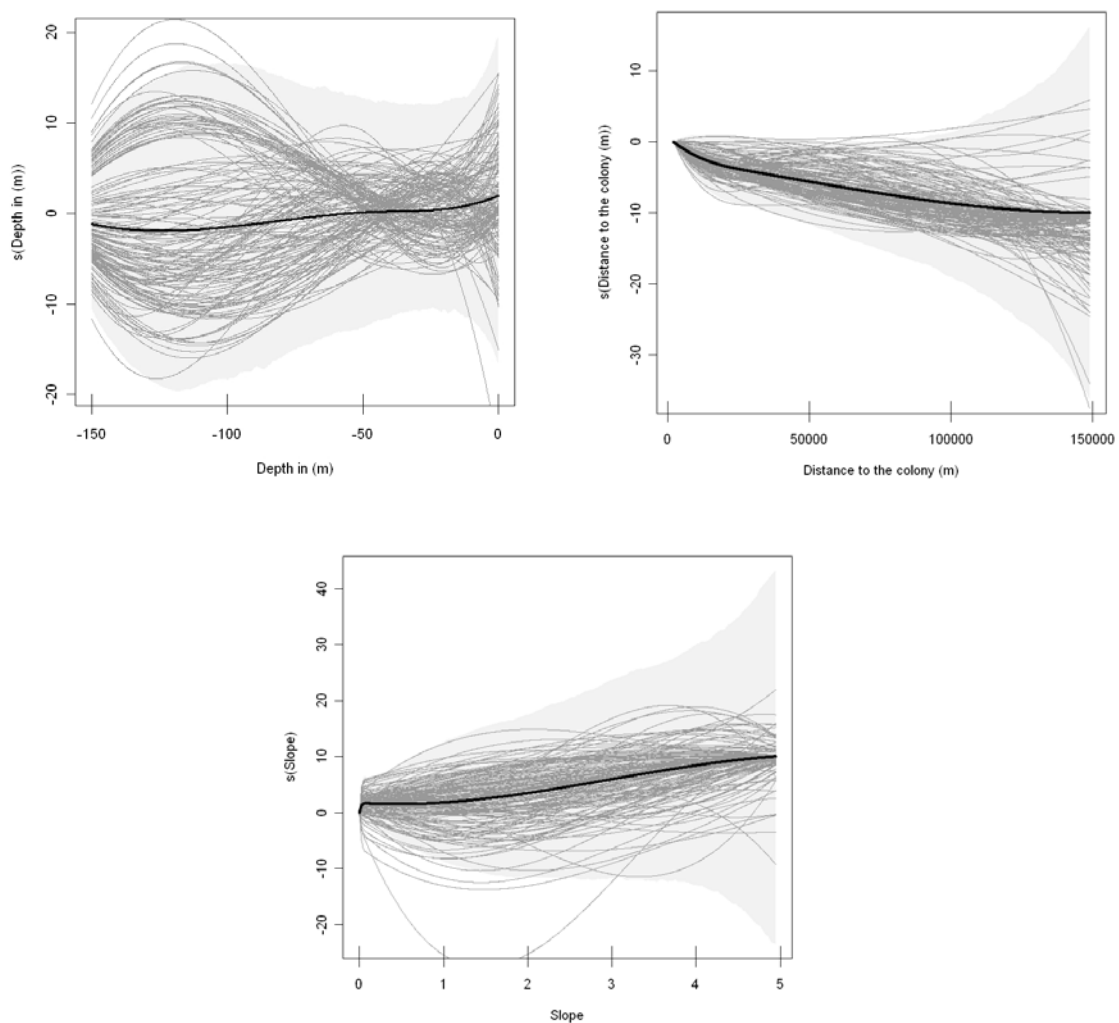
For each cell in space the habitat preference model was used to predict the likelihood that one individual would be present in that cell. Estimates of population size and the total number of days that individuals spend foraging from that site were used (see Goldsworthy and Page 2007). Using the preference model and information on foraging effort, it was possible to estimate for each cell the number of days sea lions from a particular colony spend in that cell. This calculation was applied to all cells in space and for all colonies.

## 3. Results

### 3.1 The effect of environmental variables on preference

Figure 7 shows the sea lions' preference for the three environmental variables; depth, distance to colony and slope. These figures show in the first place that there is large individual variability in preference, especially for depth. All variables are significant ( $p < 0.01$ ), but these p-values are probably underestimated, due to residual spatial and temporal correlation (Aarts et al. 2008). Due to the large sample size it is computationally not possible to incorporate such correlations in errors into the model.

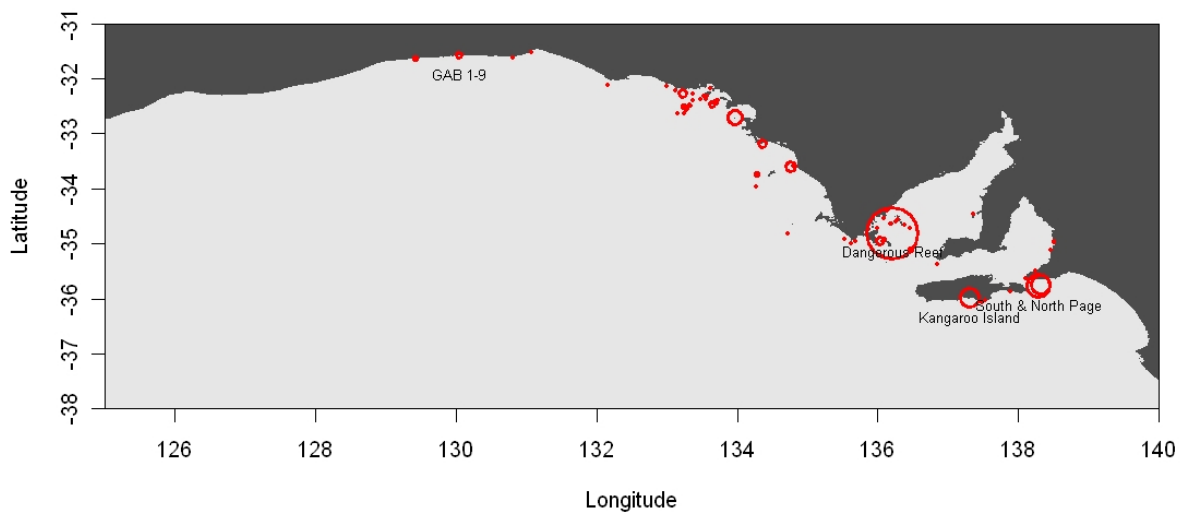
Generally sea lions avoid deep areas, areas far away from the colony and prefer regions with a steeper slope.



**Figure 7.** Sea lion preference for environmental variables: depth, distance to colony and slope. The x-axis denotes the explanatory variable. The y-axis is a measure of preference. A higher value on the y-axis indicates higher preference. The black line represents the mean population response. The grey lines individual specific random effects and the grey shaded area the 95% confidence interval. The figure shows that preference decreases with depth, decreases with distance to colony and increases with steeper slopes.

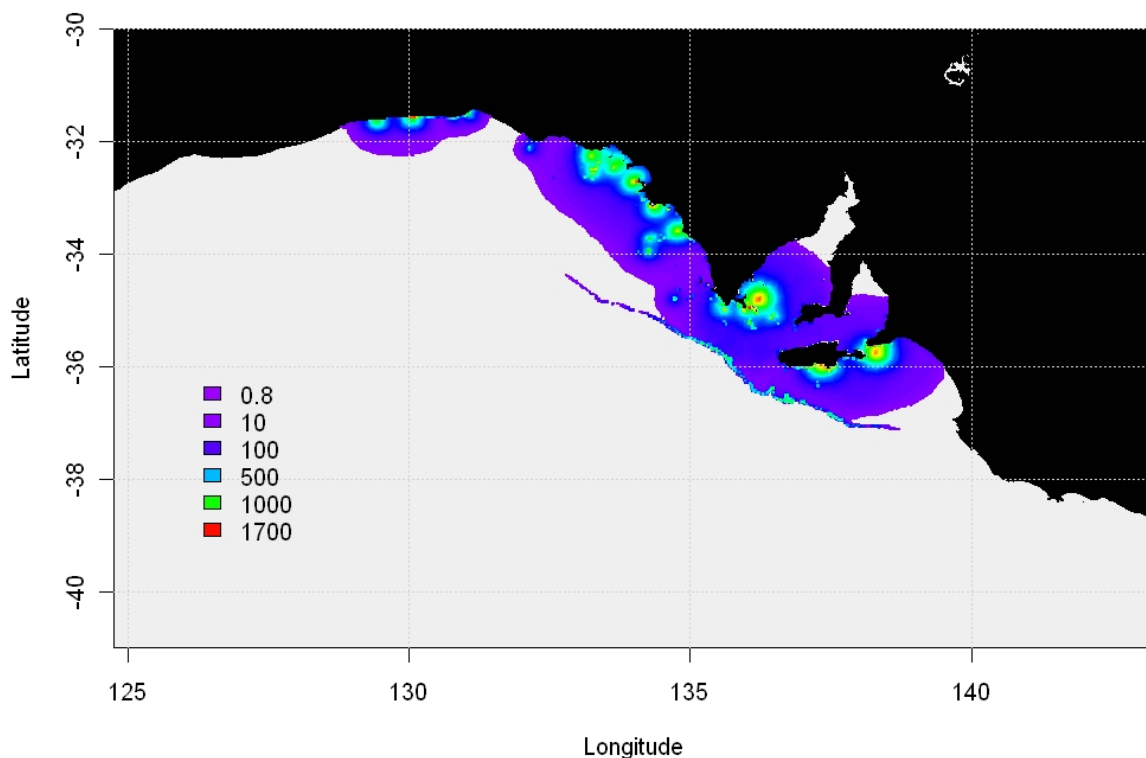
### 3.2 Spatial prediction of usage

The spatial distribution of Australian sea lions is influenced by both the habitat preference model and the distribution of foraging effort among the different colonies. Fig. 8 shows the female forage effort for the different colonies. The largest colony is 'Dangerous Reef' with approximately 200,000 female seal foraging days. Also 'North Page', 'South Page' and 'Kangaroo Island' account for approximately 28% of the total sea lion foraging days. The most western colonies (GAB B1 – B9) only account for about 5% of the foraging effort in South Australian waters.



**Figure 8.** Distribution of foraging effort (number of sea lion days foraging) associated with each colony

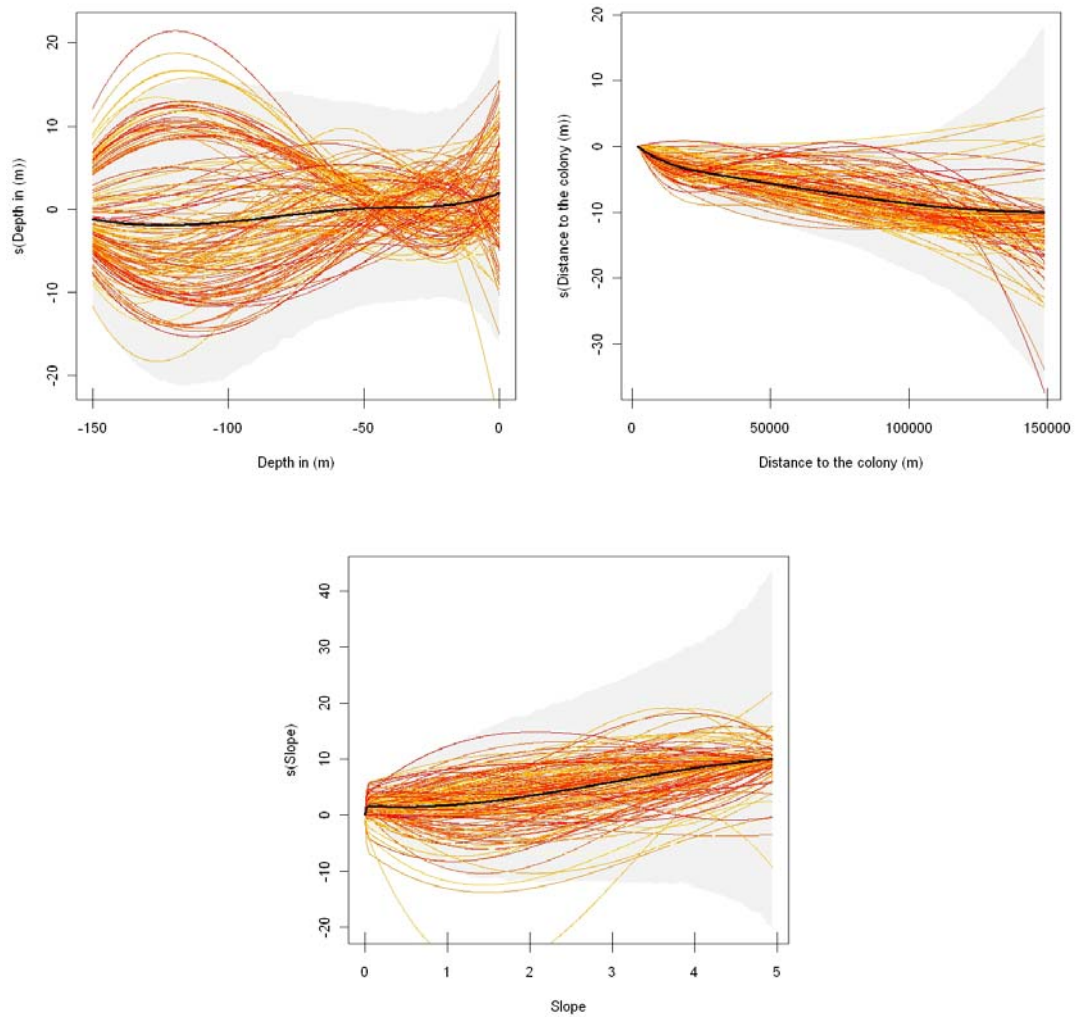
The habitat preference model (Fig. 7 and eq. 1) can be used to estimate for each individual foraging from a particular colony, the probability that it will be found in a particular cell. Because information on the total foraging effort per colony (see Fig. 8) is known, it is possible to rescale these predictions to predict the distribution of the entire colony. Fig. 9 shows the predicted distribution of the cells in space that contain 95% of the total foraging effort.



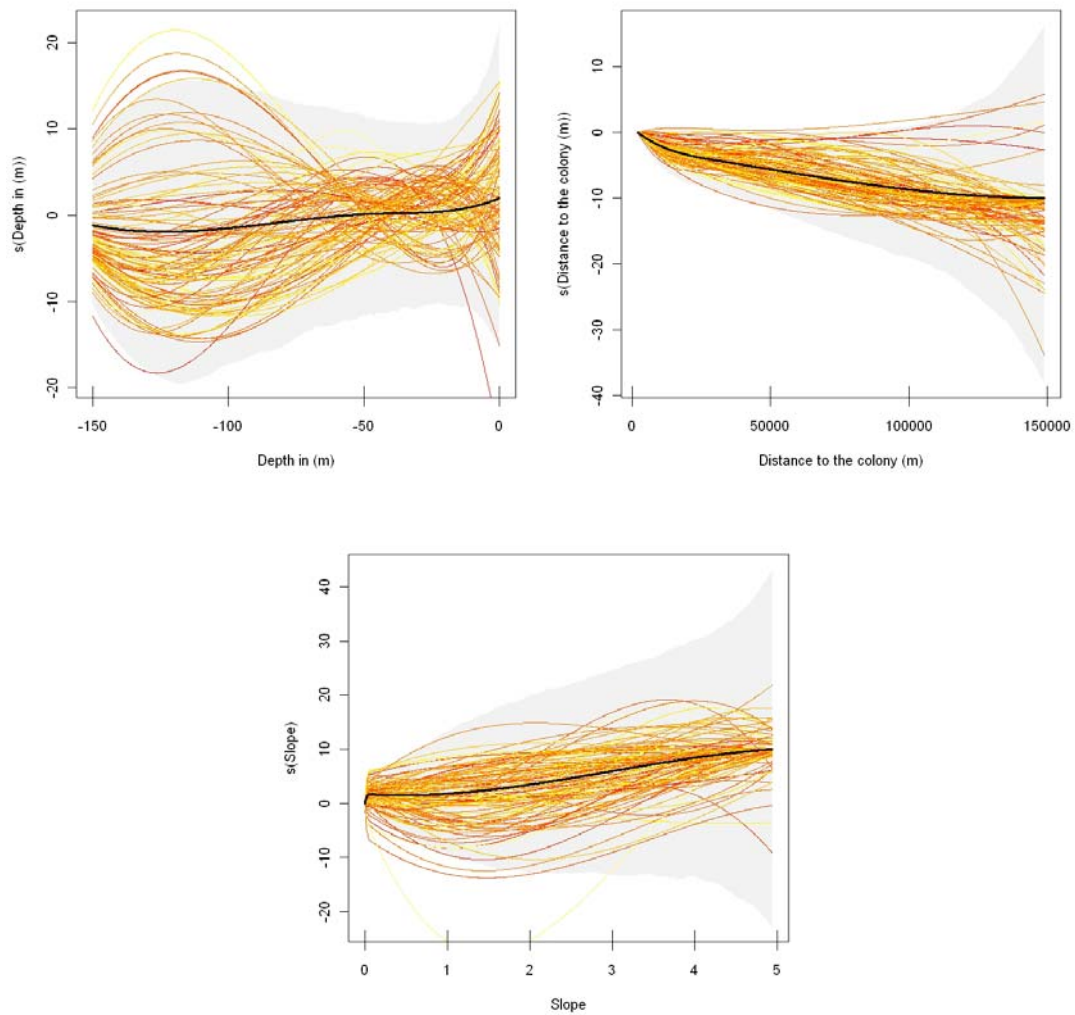
**Figure 9.** The spatial prediction of usage at sea. The colours represent the total seal days at-sea for each 0.01 by 0.01 ° cell. The coloured area contains 95% core area of there distribution at sea.

### 3.3 Individual difference in habitat preference

Each grey line in Fig. 7 represents one individual. Individuals differ from each other, e.g. in where it was caught, age and size. Such individual difference may also reflect individual preference for habitats. Therefore, we also visually inspected if some of the individual variability may have been influenced by where the animal was caught (described by its longitudinal position) and the size of the individual (expressed by its length). Due to the lack of information on the mass of the animal only the effect of length was investigated. These results are represented in Fig. 10 and Fig. 11 respectively. The bold black line reflects the mean population preference for the environmental variable in question. Each coloured line reflects the preference of one individually tracked animal. Any visible pattern in the colours, such as clusters of lines with similar colours or colour gradients, indicate difference in preference that may be explained by those individual characteristics. Fig. 10 and 11 suggest that some groups of individuals seem to have similar preferences (i.e. a cluster of red lines in Fig. 10a), however this is not explained by the longitudinal position or length of the animal.



**Figure 10** Sea lion preference for environmental variables: depth, distance to colony and slope (see legend Fig. 7), and the difference in individual preference for the environmental variables colour coded by longitudinal position. Red indicates sea lions caught at high longitudes and yellow at low longitudes



**Figure 11** Sea lion preference for environmental variables: depth, distance to colony and slope (see legend Fig. 7), and the difference in individual preference for the environmental variables colour coded by length. Red indicates long (i.e. large) sea lions and yellow short (i.e. small) individuals.

## 4. Discussion

The major objective of this study was to conduct a first exploratory analysis (concentrating on adult females) to get an idea about the major explanatory variables influencing the spatial distribution at sea. Secondly, we investigated whether it is possible to make spatial predictions at the population level. In this modelling exercise we only included distance to the colony, slope and depth. Not all individual specific effects were included and no extensive model selection and validation using test data has been incorporated. Nevertheless, this work should provide a solid guideline for more in-depth further analysis.

### 4.1 Spatial distribution and habitat preference of the Australian sea lion

The spatial distribution of Australian sea lions is reasonably well predicted by the model. It is however largely influenced by one and perhaps the most obvious explanatory variable: distance to colony (Fig. 7 & 9). Visually comparing the model predictions and actual distribution of satellite telemetry locations, suggest that in the eastern region, the model seems to predict the animal observations well. Also the relative higher usage at the edge of the continental slope is predicted by the model – although not supported by individual tracking data. In contrast, the prediction for the western colonies does not match well with the telemetry observations. A large number of animal locations from these colonies are outside the 95% core area, which may suggest that animals spend more time further away from the colony than predicted by the model. This miss-match may have two reasons. First of all, the forage effort (the number of days sea lions spend away, *i.e.* foraging when departing from this colony) is relative small compared to some other colonies. The total forage effort from the colonies GAB B1 to B9 is only 5% of the total sea lion foraging effort. Therefore, the large number of telemetry observations are a consequence of a relatively high tagging effort in these regions (compared to the population size), and the prediction may be a reasonable accurate depiction of the distribution of the total population. A second reason may be that the individuals in these colonies behave differently. Indeed, according to Fig. 10 at least 5 western individuals (yellow lines) have a relative higher preference for areas further away from the colony.

This study reveals that on average sea lions avoid deep areas. If sea lions forage close to the bottom, increasing depths may limit the time animals can spend foraging at the bottom and this may explain this relationship. However, this does not seem to hold for all individuals (see also Fig. 7). Another interesting aspect is the sea lions' preference for steeper slopes. This preference is reflected by a large majority of individuals, also by those individuals that do not spend time at or near to the edge of the continental shelf. Therefore it could be questioned whether this preference is real or is it a consequence of colinearity with other environmental variables not included in the model. Perhaps these areas are characterized by higher primary and secondary productivity, or perhaps it is the consequence of colinearity with another variable (e.g. sediment type) not included in the model.

### 4.2 Incorporating the right environmental variables

The two most appealing aspect of modelling the spatial distribution in relation to environmental variables is that it provides an insight into which habitats sea lions prefer and that this knowledge can be used to predict the distribution for other regions where colonies are present. Sampling effort is rarely balanced across the different colonies and is almost never proportional to the population size at these colonies. This is however a requirement to use satellite telemetry data directly to estimate the distribution of the population. The ability to predict the spatial distribution is essential when no individuals have been tracked at colonies or even entire regions of space, provided that there is data on their distribution at colonies on land. Having at least some estimates about their spatial distribution is essential for managing the species and interactions with fisheries.

A downside is that for those areas where extensive tagging has taken place, the model prediction may be a poorer representation of their actual distribution. The quality of such predictions heavily depends on the availability of appropriate environmental variables and our ability to include other important aspects (e.g. intrinsic population processes such as density dependent effects between and within species) into the model.



To better understand why individuals visit certain places it essential to improve our understanding about the choices individuals make and to have data on all the relevant environmental variables. Future work in progress, such as fine – scale habitat preference studies using e.g. GPS transmitters, in stead of coarser ARGOS tags, may shed more light on this. Also first attempts are being made to film the foraging behaviour of Australian sea lions using animal-borne cameras. Both sources of information should provide a better understanding of why animals forage at certain places. This knowledge could provide a future basis for more in-depth habitat preference studies.

## Acknowledgements

We thank Page Bradley for all data preparation in a busy time when he had other important things on his mind. We thank Peter Reijnders for all his useful comments and suggestions.

## References

- Aarts, G., M. MacKenzie, B. McConnell, M. Fedak, and J. Matthiopoulos. 2008. Estimating space-use and habitat preference from wildlife telemetry data. *Ecography* **31**:140-160.
- Boyce, M. S., and L. L. McDonald. 1999. Relating populations to habitats using resource selection functions. *Trends in Ecology & Evolution* **14**:268-272.
- Goldsworthy, S. D., and B. Page. 2007. A risk-assessment approach to evaluating the significance of seal bycatch in two Australian fisheries. *Biological Conservation* **139**:269-285.
- Guisan, A., and N. E. Zimmermann. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* **135**:147-186.
- Hastie, T., and R. J. Tibshirani. 1990. *Generalized Additive Models*. Chapman and Hall, London.
- Johnson, C. J., S. E. Nielsen, E. H. Merrill, T. L. McDonald, and M. S. Boyce. 2006. Resource selection functions based on use-availability data: Theoretical motivation and evaluation methods. *Journal of Wildlife Management* **70**:347-357.
- Keating, K. A., and S. Cherry. 2004. Use and interpretation of logistic regression in habitat selection studies. *Journal of Wildlife Management* **68**:774-789.
- Lele, S. R., and J. L. Keim. 2006. Weighted distributions and estimation of resource selection probability functions. *Ecology* **87**:3021-3028.
- Matthiopoulos, J. 2003. The use of space by animals as a function of accessibility and preference. *Ecological Modelling* **159**:239-268.
- Matthiopoulos, J., B. McConnell, C. Duck, and M. Fedak. 2004. Using satellite telemetry and aerial counts to estimate space use by grey seals around the British Isles. *Journal of Applied Ecology* **41**:476-491.
- Mysterud, A., and R. A. Ims. 1998. Functional responses in habitat use: Availability influences relative use in trade-off situations. *Ecology* **79**:1435-1441.
- Wood, S. 2006. *Generalized Additive Models: An introduction with R*. Chapman & Hall, New York.

## Referees and Authors

Rapport C107/08

Project Number: 4396111201

This report has been professionally prepared by Wageningen IMARES. The scientific validity of this report has been internally tested and verified by another researcher and evaluated by the Scientific Team at Wageningen IMARES.

Approved: Peter Reijnders  
Senior Researcher

Signature:



Date: December 2007

Approved: Drs. F.C. Groenendijk  
Head of department

Signature:



Date: December 2008

Number of copies: 12  
Number of pages: 18  
Number of graphs: 11